From DNA to Protein

**Transcription**
- Explain the purpose for this process and its sub-cellular compartment.
- Learn the three steps of transcription and the bio-molecules involved.
- Name the stretch of nucleotides in the DNA that binds RNA polymerase to initiate transcription.
- Name the DNA stretch that causes RNA polymerase to come off DNA terminating transcription.
- Explain the split nature of eukaryotic genes. Distinguish between exon, intron. Which contains information that will specify the amino acid sequence of the protein product?
- Explain RNA processing in human cells.

**Genetic code**
- Explain the language of nucleic acids: letters and words (nucleotides and codons), number of nucleotides that make up a codon, and total number of codons.
- Which codon marks the initiation of translation, what amino acid does it specify? How many stop codons specify the termination of translation.

**Translation or protein synthesis**
- Where does it take place, what components are necessary, and what are the three steps.
- Which biomolecule can interpret the language of nucleic acids into the language of proteins.
- Recognize the importance of proper protein folding for functional activity of the protein and the roles of chaperones and proteasomes.
Genetic Basis of Development

From a diploid zygote to a multi-cellular organism

Sperm cell

Egg cell

Fertilized egg with DNA from both parents

Embryo’s cells with copies of inherited DNA

Offspring with traits inherited from both parents
Gene Expression

The human genome contains about 20,325 protein-encoding genes
- However, this represents only a small part of the genome

Much of the human genome *controls* protein synthesis
- Including the time, speed, and location
THE FLOW OF GENETIC INFORMATION
The DNA of the gene is transcribed into RNA which is translated into the polypeptide (protein)
Proteins have diverse functions in the body

<table>
<thead>
<tr>
<th>Protein</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actin, myosin, dystrophin</td>
<td>Muscle contraction</td>
</tr>
<tr>
<td>Antibodies, antigens, cytokines</td>
<td>Immunity</td>
</tr>
<tr>
<td>Carbohydrases, lipases, proteases, nucleases</td>
<td>Digestion (digestive enzymes)</td>
</tr>
<tr>
<td>Casein</td>
<td>Milk protein</td>
</tr>
<tr>
<td>Collagen, elastin, fibrillin</td>
<td>Connective tissue</td>
</tr>
<tr>
<td>Colony-stimulating factors, erythropoietin</td>
<td>Blood cell formation</td>
</tr>
<tr>
<td>DNA and RNA polymerase</td>
<td>DNA replication, gene expression</td>
</tr>
<tr>
<td>Ferritin</td>
<td>Iron transport in blood</td>
</tr>
<tr>
<td>Fibrin, thrombin</td>
<td>Blood clotting</td>
</tr>
<tr>
<td>Growth factors, kinases, cyclins</td>
<td>Cell division</td>
</tr>
<tr>
<td>Hemoglobin, myoglobin</td>
<td>Oxygen transport</td>
</tr>
<tr>
<td>Insulin, glucagon</td>
<td>Control of blood glucose level</td>
</tr>
<tr>
<td>Keratin</td>
<td>Hair structure</td>
</tr>
<tr>
<td>Tubulin, actin</td>
<td>Cell movements</td>
</tr>
<tr>
<td>Tumor suppressors</td>
<td>Cancer prevention</td>
</tr>
</tbody>
</table>
Gene Expression

Production of protein from instructions on the DNA

Gene expression requires several steps

- **Transcription** = Production of mRNA
- **Translation** = Production of protein using mRNA, tRNA, and rRNA
- Folding of the protein into the active 3-D form
The directional flow of genetic information

The DNA is transcribed into RNA which is translated into the polypeptide (protein)

Figure 10.1
Transcription

RNA is the bridge between DNA and protein

RNA is synthesized from one strand of the DNA double helix, which is called the template strand

The complementary strand is called the coding strand of DNA

Requires the enzyme RNA polymerase
Transcription

Figure 10.2
Nucleic Acids

There are two types of nucleic acids
- RNA
- DNA

Both consist of sequences of N-containing bases joined by sugar-phosphate backbones
- However, they differ in several aspects
## Nucleic Acids

### Table 10.2  How DNA and RNA Differ

<table>
<thead>
<tr>
<th>DNA</th>
<th>RNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Usually double-stranded</td>
<td>1. Usually single-stranded</td>
</tr>
<tr>
<td>2. Thymine as a base</td>
<td>2. Uracil as a base</td>
</tr>
<tr>
<td>3. Deoxyribose as the sugar</td>
<td>3. Ribose as the sugar</td>
</tr>
<tr>
<td>4. Maintains protein-encoding information</td>
<td>4. Carries protein-encoding information and controls how information is used</td>
</tr>
<tr>
<td>5. Cannot function as an enzyme</td>
<td>5. Can function as an enzyme</td>
</tr>
<tr>
<td>6. Persists</td>
<td>6. Transient</td>
</tr>
</tbody>
</table>
DNA
Stores RNA- and protein-encoding information, and transfers information to daughter cells

RNA
Carries protein-encoding information, and helps to make proteins

Double-stranded

Generally single-stranded

Deoxyribose as the sugar

Ribose as the sugar

Bases used:

Thymine (T)
Cytosine (C)
Adenine (A)
Guanine (G)

Bases used:

Uracil (U)
Cytosine (C)
Adenine (A)
Guanine (G)
Types of RNA

There are three major types of RNA
- messenger RNA or mRNA
- ribosomal RNA or rRNA
- transfer RNA or tRNA

Other classes of RNA control gene expression
- Will be discussed in Chapter 11
## Major Types of RNA

<table>
<thead>
<tr>
<th>Type of RNA</th>
<th>Size (number of nucleotides)</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>mRNA</td>
<td>500–4,500 +</td>
<td>Encodes amino acid sequence</td>
</tr>
<tr>
<td>rRNA</td>
<td>100–3,000</td>
<td>Associates with proteins to form ribosomes, which structurally support and catalyze protein synthesis</td>
</tr>
<tr>
<td>tRNA</td>
<td>75–80</td>
<td>Transports specific amino acids to the ribosome for protein synthesis</td>
</tr>
</tbody>
</table>
mRNA

Carries information that specifies a particular protein

Produced in the nucleus

Transported to the ribosome

A three nucleotide **codon** specifies a particular amino acid

Most mRNAs are 500-4,500 bases long
rRNA

Associate with proteins to make up ribosomes
Ribosomes consist of two subunits that join during protein synthesis
rRNAs provide structural support
- Some are a catalyst (ribozymes)
Most rRNAs are from 100-3,000 bases long
rRNA are components of the ribosome

Figure 10.4
tRNA

Only 75-80 bases long
The 2-D shape is a cloverleaf shape
The 3-D shape is an inverted L
Has two business ends:
- The **anticodon** forms hydrogen bonds with the mRNA codon
- The 3’ end binds the amino acid specified by the mRNA codon
Steps of Transcription

Transcription is described in three steps:

- Initiation
- Elongation
- Termination
Steps of Transcription

In transcription initiation, a cascade of transcription factors bind to the promoter region of a gene.

These open a pocket allowing the RNA polymerase to bind just in front of the start of the gene sequence.
Transcription Factors

In more complex organisms transcription factors control gene expression and link genome to environment

- These contain DNA-binding domains

About 2,000 in humans

Mutations in transcription factors may cause a wide range of effects
Figure 10.7

DNA

Promoter

Gene sequence to be transcribed

3’ GG TATA CCC 5’

a. TATA box

TATA binding protein

3’ GG TATA CCC 5’

b. Transcription factor

RNA polymerase

3’ GG TATA CCC 5’

c. Transcription begins
Steps of Transcription

During elongation, RNA polymerase reads the nucleotides on the template strand from 3’ to 5’ and creates an RNA molecule that looks like the coding strand.

Then termination occurs when sequences in the DNA prompt the RNA polymerase to fall off ending the transcript.
Simultaneous Transcription of mRNAs

Several mRNAs may be transcribed from the same template DNA strand at a time.
eukaryotes

Pre-mRNA undergoes processing to remove introns before mRNA leaves the nucleus
Figure 10.10

Split genes

Processing of pre-mRNA in eukaryotes
http://highered.mcgraw-hill.com/sites/0072437316/student_view0/chapter15/animations.html#
RNA Processing

In eukaryotes, mRNA transcripts are modified before they leave the nucleus. Several steps process pre-mRNA into mature mRNA:

1) A methylated cap is added to the 5’ end
   - Recognition site for protein synthesis
2) A poly-A tail is added to the 3’ end
   - Stabilizes the mRNA
3) Splicing occurs
- Introns (“intervening sequences”) are removed
- Exons (“expressed sequences”) are spliced together
- Note that introns may outnumber and outsize exons

Finally, the mature mRNA is sent out of the nucleus
The genetic information is coded and is translated into amino acid sequences

The “words” of the DNA “language” are triplets of bases (3 bases long) called codons

Each codons in a gene specify one amino acid sequence of the polypeptide
Translation

The process of reading the mRNA base sequence and creating the amino acid sequence of a protein

Occurs on the ribosome

Figure 10.11
The Genetic Code

It is a triplet code
- Three successive mRNA bases form a codon

There are 64 codons, including:
- One start signal (AUG)
- Three stop signals (UAA, UAG, and UGA)
Figure 10.12
In-class activity/Genetic code

Use the genetic code table to answer the following questions:

1. How many codons are there for leu (leucine)?
2. How many codons are there for Met (Methionine)?
3. How many codons are there for Phe (phenylalanine)?

Draw a conclusion about the number of codons for amino acids.

4. How many “stop” codons are there?

Answer the following questions using this genetic code: 5’-AUGACCCCUUUGUUAUACUAA-3’

5. How long is this message in nucleotides?

6. Is this the information present in DNA or in mRNA? Explain your answer.

7. Write down the sequence of amino acids coded for by the above stretch of nucleotides.

how long is this polypeptide?
The genetic code is the Rosetta stone of life

All organisms use the same genetic code
The Genetic Code

It is non-overlapping

It is degenerate
- Two or more codons may specify the same amino acid (*synonymous codons*)

It is universal
- Evidence that all life evolved from a common ancestor
Reading Frame

A sequence of amino acids encoded from a certain starting point in a DNA/RNA sequence

Figure 10.14
Mutation (Learning Objectives)

Define the term mutation and its sources

Explain the difference between the general types of mutations: base substitutions or point mutations, insertions, and deletions.

Explain the three types of point mutations: silent, mis-sense, and nonsense.

Understand what is meant by the reading frame and the effect on insertions or deletions.

Apply and demonstrate your understanding of mutations, the genetic code, and the flow of genetic information in the cell, by predicting the effect of a particular mutation on the final protein product.
Genetic Mutation

A change in nucleotide sequence of DNA
Occurs spontaneously
Can be induced by chemical and physical agents
**Mutation/ In-class**

1. For amino acids with redundant codons, which nucleotide position(s) are always the same, i.e. conserved? (marked next to the genetic code table)

2. Which amino acid does UUA code for?

3. Does a mutation that changes the codon UUA into a UUG change the amino acid sequence at the protein level?

4. Does a DNA mutation changing the codon UUA into a UCA change the amino acid sequence at the protein level?

5. What impact would a change of the codon UUA into a UAA have at the translational level?
Types of Mutations

1. **Base substitution** (point mutation)

2. **Insertions and deletions**
Types of Mutations

1. **Base substitution** (point mutation)
   
a. **missense** replaces one amino acid with another.

   b. **nonsense** changes codon into a stop codon (results in truncated or shorter protein)

   c. **silent** changes replaces one redundant codon with another (mutation at DNA level with no effect on the amino acid sequence of the protein)
Types of Mutations

1. **Base substitution** (point mutation)
   a. **missense** - replaces one amino acid with another.
   b. **nonsense** changes codon into a stop codon (results in truncated or shorter protein)
   c. **silent** changes replaces one redundant codon with another (mutation at DNA level with no effect on the amino acid sequence of the protein)

2. **Insertions and deletions**
   Result in frame-shifts
Reading Frames

mRNA sequence: A U G G C A U U G C C C U U A A U

Reading frame #1:
- A U G
  - Methionine
- G C A
  - Alanine
- U G
  - Leucine
- C C U
  - Proline
- U A U
  - Tyrosine

Reading frame #2:
- A U G G
  - Tryptophan
- C A U
  - Histidine
- U G C
  - Cysteine
- C U U
  - Leucine

Reading frame #3:
- A U G G C
  - Glycine
- A U U
  - Isoleucine
- G C C
  - Alanine
- U U A
  - Leucine
Substituting, inserting, or deleting nucleotides alters a gene with varying effects on the organism.
Translation

Requires mRNA, tRNAs with amino acids, ribosomes, energy molecules (ATP, GTP) and protein factors

Divided into three steps:
- Initiation
- Elongation
- Termination
Translation Initiation

The leader sequence of the mRNA forms H-bonds with the small ribosomal subunit

The start codon (AUG) attracts an initiator tRNA that carries methionine

This completes the initiation complex
Figure 10.15

Translation initiation

DNA → Replication → RNA → Transcription → Translation → Protein

Nucleus → mRNA → Small ribosomal subunit

mRNA: 5' UUCGUCACUGGGAUGUAAGCGAAA 3'

Leader sequence

Assembling to begin translation

Protein factors + energy source

Initiator tRNA: Met

Initiation complex

Small ribosomal subunit

mRNA: 5' UUCGUCACUGGGAUGUAAGCGAAA 3'
Translation Elongation

The large ribosomal subunit joins
A second tRNA binds to the next mRNA codon
First peptide bond forms between the two amino acids
  - Catalyzed by an rRNA ribozyme

tRNAs bring in more amino acids, as the ribosome moves down the mRNA
a. **Second amino acid joins initiation complex.**

b. **First peptide bond forms as new amino acid arrives.**

c. **Amino acid chain extends.**
Translation Termination

Occurs when a stop codon enters the A site of the ribosome

A protein release factor frees the polypeptide

The ribosomal subunits separate and are recycled
Figure 10.17

a. Ribosome reaches stop codon.

b. Components disassemble.

http://vcell.ndsu.edu/animations/translation/movie-flash.htm
Multiple Copies of a Protein Can be Made Simultaneously

The closer to the end of the gene, the longer the polypeptide
<table>
<thead>
<tr>
<th>Type of Molecule</th>
<th>Rules and Relationships</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNA coding strand</td>
<td>1  Coding and template strands have complementary DNA bases.</td>
</tr>
<tr>
<td></td>
<td>2  mRNA is complement of DNA template strand, with U for T.</td>
</tr>
<tr>
<td>DNA template strand</td>
<td>3  mRNA is same as DNA coding strand, with U for T.</td>
</tr>
<tr>
<td></td>
<td>4  tRNA anticodons are complement of mRNA.</td>
</tr>
<tr>
<td>mRNA codons</td>
<td>5  tRNA anticodons are same as DNA template strand, with U for T.</td>
</tr>
<tr>
<td></td>
<td>6  tRNA anticodons are complement of DNA coding strand, with U for T.</td>
</tr>
<tr>
<td>tRNA anticodons</td>
<td>7  tRNA translates genetic code, bringing together amino acids specified by DNA coding strand.</td>
</tr>
<tr>
<td>Amino acids (protein)</td>
<td>8  Amino acids bond to form a protein.</td>
</tr>
</tbody>
</table>
Protein Structure

Proteins fold into one or more 3-D shapes or conformations
Protein Folding

Protein folding begins as translation proceeds

Enzymes and *chaperone proteins* assist

Should a protein misfold, an “unfolded protein response” occurs
- Protein synthesis slows or even stops
Protein Misfolding

Misfolded proteins are tagged with ubiquitin. Then, they are escorted to a **proteasome**, a tunnel-like multiprotein structure. As the protein moves through the tunnel, it is straightened and dismantled.

Proteasomes also destroy properly-folded proteins that are in excess or no longer needed.
Protein Misfolding

Copyright © The McGraw-Hill Companies, Inc. Permission required for reproduction or display.

Ubiquitin Conjugation

Protein Degradation

Protein → Ubiquitin molecules → Proteasome → Peptides → Amino acids

Figure 10.20

http://bioisolutions.blogspot.com/2007/05/proteasomes.html
# Protein Misfolding

## Table 10.6 Disorders Associated with Protein Misfolding

<table>
<thead>
<tr>
<th>Disease</th>
<th>Misfolded Protein</th>
<th>MIM (protein)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alzheimer disease</td>
<td>Amyloid beta precursor protein</td>
<td>104760</td>
</tr>
<tr>
<td>Familial amyotrophic lateral sclerosis</td>
<td>Superoxide dismutase</td>
<td>147450</td>
</tr>
<tr>
<td>Huntington disease</td>
<td>Huntingtin</td>
<td>143100</td>
</tr>
<tr>
<td>Parkinson disease</td>
<td>Alpha synuclein</td>
<td>163890</td>
</tr>
<tr>
<td>Lewy body dementia</td>
<td>Alpha synuclein</td>
<td></td>
</tr>
<tr>
<td>PKU</td>
<td>Phenylalanine hydroxylase</td>
<td>261600</td>
</tr>
<tr>
<td>Prion disorders</td>
<td>Prion protein</td>
<td>176640</td>
</tr>
</tbody>
</table>

(All but Huntington disease are genetically heterogeneous; that is, abnormalities in different proteins cause similar syndromes.)

## Examples of Familial Prion Disorders
- Creutzfeldt-Jakob Disease (CJD) in humans
- Fatal Familial Insomnia (FFI)
Prions

Prion protein (PrP) can fold into any of several conformations

One conformation is aberrant

- Moreover, it can be passed on to other prions upon contact, propagating like an “infectious” agent

In addition, the aberrant conformation can form even in the wild-type protein
Prions

Figure 10.22

[Image of a neuron showing the nucleus, cytoplasm, and intracellular membrane. Adjoining boxes illustrate the conversion of normal PrP to scrapie PrP.]

Copyright © The McGraw-Hill Companies, Inc. Permission required for reproduction or display.